





The NCCS has been a leading capacity computing facility, providing a production environment and support resources to address the challenges facing the Earth and space sciences research community.

Hardware resources

The year 2000 marks a transition as the NCCS prepares to acquire next-generation systems that will significantly enhance the computing capacity. The NCCS mass storage capacity has continued to increase rapidly as we have migrated to denser and faster media.

Computing

The NCCS primarily provides two types of computing architectures: vector and parallel supercomputing.

Vector supercomputing allows a single computer to process a large amount of data (vector) with a

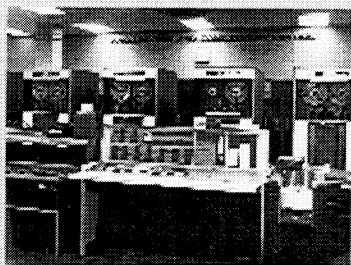
single instruction. The NCCS is committed to maintaining and supplying some of the fastest vector machines possible for its users. The NCCS maintains a pair of 24-processor Cray SV1 supercomputers, each of which can perform up to 24 billion floating-point operations per second (GigaFLOPS). Most of the smaller research projects that use NCCS resources are assigned to these systems. Long-standing software that has not been modified to take advantage of parallel computing is best suited for these computers, which provide large shared-memories. However, the SV1 systems provide for modest processing capability.

Specifications: Cray SV1 systems

- 8 gigabytes (GB) of main memory
- 7.6 terabytes (TB) of compressed disk storage (5 TB uncompressed)
- 24 GigaFLOPS
- A pair of StorageTek (STK) Automated Cartridge System tape drive silos

HISTORY

1960s and 1970s



Goddard has a long history of providing the most powerful computing available to its science community. This computing power began with the introduction of Goddard's first scientific processor, an IBM 7090, in 1960. The IBM 7090 was followed by its successor, the IBM 360/91, in the late 1960s. As the computer hardware evolved, so did the operating environment. Terminals ushered in a shift from boxes of cards to batch file systems.



- Eight STK Timberline 9490 cartridge tape drives
- A Powderhorn Robot and a Wolfcreek Robot
- UNICOS 10.0.0.7 operating system
- Compilers for C, C++, and Fortran 90
- Libraries: BLAS, EISPACK, FISH-PACK, HDF, IMSL, NAG, ODEPACK, and SLATEC
- Utilities: FLINT, prof, perftrace, hpm, atexpert, and TotalView

Parallel computing architectures connect multiple processors within a single computer system. These processors collectively work on the same calculating task. The NCCS maintains a parallel-processing SGI Origin 2000, which primarily supports research at the Data Assimilation Office (DAO). The Origin 2000 allows programs to use all 64 processors and the entire memory to work on a single problem. The memory appears similar to the shared memory of the Cray.

Specifications: SGI Origin 2000

- 64 R12000 processors
- Cache-coherent Non Uniform Memory Access (ccNUMA) architecture
- 32 GB of main memory
- 1,559 GB of disk storage
- SGI Irix 6.5 operating system
- Compilers for C, C++, Fortran 77, and Fortran 90

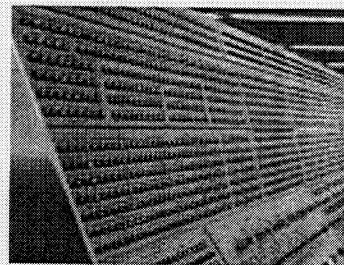
In addition, the NCCS operates a large Cray T3E. The T3E is one of the world's most powerful parallel supercomputers. It can perform a staggering 778 billion floating-point calculations in a single second. Software that is designed to take full advantage of many processors and the short-term memory storage of a computer's cache can perform very well on the T3E.

The T3E is the workhorse of the NASA Seasonal-to-Interannual Prediction Project (NSIPP). The

HISTORY

Early 1980s

An IBM 3081 replaced the IBM 360/91. The interactive processing capabilities of the IBM 3081K literally added a new dimension to scientific investigations. By providing a full range of graphics tools and the means to access large databases, the computing center enabled scientists to analyze their data at individual terminals and workstations. Although a more convenient method for the scientists, this new technology required storage for larger volumes of on-line data.





NSIPP uses 1,000 of the T3E's 1,360 processors. The Grand Challenge Investigations program of NASA's Earth and Space Sciences Project uses the remaining processors.

Specifications: Cray T3E

- 1,360 processors
- 778 GigaFLOPS
- 170 GB of main memory
- 2,205 GB of disk storage

Mass data storage and delivery

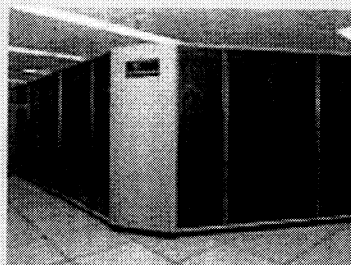
Mass storage is the natural complement to performance computing because the output from some computer model simulations can be as large as a terabyte (TB), that is, 1 trillion bytes. Output of that size is routinely broken down into more manageable units for ease of retrieval, creating even more files.

Accommodating today's growing data storage needs is a top priority of the NCCS. Never before has the science world produced such massive quantities of data. In FY 2000, more than 6 million files containing more than 150 TB of user data are under the control of the NCCS mass storage system. The mass storage system operates under a Sun E10000 server. Information stored on this system is organized by the UniTree Central File Manager.

UniTree is an intelligent hierarchical data archival system. It assigns data to one of several levels of storage media, depending on how often a particular piece of information is required. Initially, all data is recorded onto both magnetic disk and tape cartridge storage. Because of its quick loading speed, magnetic disk space is prioritized for information that is frequently accessed or updated. UniTree's disk cache has been expanded more than three times over, from 1.5 TB to 5 TB capacity. When space is needed on the magnetic disks, UniTree automatically erases information that has not been used recently. The information remains intact on tape cartridge,

HISTORY

Mid to late 1980s



With the formation of a new directorate focusing on research in the Earth sciences, Goddard acquired a Control Data Corporation CYBER 205 vector supercomputer. This system supported more sophisticated, computationally intensive global models to simulate physical processes more realistically than ever before. The acquisition of the CYBER 205 marked a major shift from scalar computing into vector supercomputing. The center augmented the CYBER 205 with an ETA10 supercomputer in 1988.



from which it can be retrieved when needed. To reduce the risk of data loss from tape cartridge failure, the NCCS provides duplicate data storage at a remote location.

Specifications: Sun E10000 server

- Disk cache of 5 TB
- 8 GB of memory
- 12 GB of system disk storage
- UniTree Central File Manager

Specifications: Tape Libraries

- 7 STK 9310 Powderhorn robotic storage silos
- An IBM 3494 robotic tape library

NCCS support services

The NCCS remains committed to staffing its support groups with highly qualified personnel who have backgrounds in information technol-

ogy and scientific computing. During FY 2000, the NCCS made a significant effort to enhance the services it provides to the scientific user community through user support, applications support, and systems support. The NCCS changed contractors, reorganized the support services, and increased the number of professionals experienced in high-end computing, numerical modeling, and system administration. The NCCS also extended the hours that customer service is available to users.

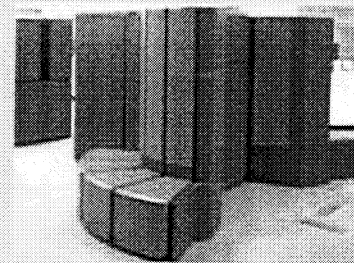
User support

The User Support Group functions include the user administration, technical assistance, and system accounting. The primary functions of this group are establishing user accounts on the NCCS systems, providing general support on effective use of the systems, and tracking system utilization by the user community. User Support Group members work closely with the application and system support groups to respond to user requirements and problems.

Acquisition of the Cray Research Cray Y-MP marked the next major increase in computational capability and the first time that direct access to the supercomputer was available. As computational power increased, the need for mass storage increased even more rapidly. This storage requirement led to the implementation of a separate mass storage subsystem.

HISTORY

Early 1990s





NCCS users can contact the Technical Assistance Group (TAG) through e-mail, telephone, phone mail, and the World Wide Web (WWW) to send questions and problem reports to the Help Desk. In FY 2000, the TAG received more than 650 requests for help-desk support. Typical questions concerned debugging and optimizing Fortran programs, using systems and applications software, solving input/output and media-related problems, backing up and restoring files, and using other installed utilities and NCCS facilities. When necessary, the TAG also develops new tools to provide the user community with improved or more efficient access to NCCS resources. Although the TAG team resolves most user issues, it can also refer user requests and reports to the applications support or system administration groups for resolution.

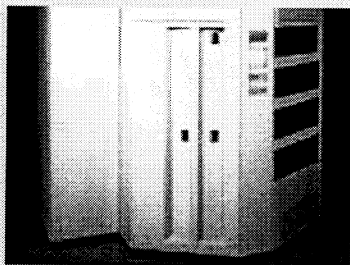
Most NCCS documentation is available to the user community through the NCCS WWW documentation server. This site provides general information about the NCCS and its resources, technical information about NCCS systems, system status information, news articles, and

minutes of Computer Users Committee (CUC) meetings. Ensuring that this information remains up-to-date and accurate is a major task for the TAG. The TAG has made significant improvements to both the navigation and content for the site and is committed to providing the best possible interface for the user community.

The NCCS User Administrator is often the first point of contact for outside users, providing assistance to NCCS division representatives, sponsors, and new users with questions, problems, comments, and suggestions. The User Administrator also processes requests from NASA Headquarters funding managers and from GSFC Directorate or Division funding personnel. Funding for a scientific research effort comes from NASA Headquarters or through a GSFC Directorate or Division, which authorizes the computing time. We are reviewing the user administration process to find ever more efficient ways to handle requests for NCCS resources, and we are beginning to develop a new web interface for adding and changing user information.

HISTORY

Early 1990s



A CONVEX mini-supercomputer and STK storage silos, each holding a TB of data, replaced the old mass storage units. This near-line storage was essential for holding the large amounts of data required for coupled land-ocean-atmosphere models.



The NCCS has also begun a total rewrite of the user accounting system to provide improved access to resource utilization information by users, sponsors, and the NCCS management. The new system will provide the potential for user and sponsor access to daily updates to the accounting information through the web, and the monthly accounting reports will be available electronically earlier than through the current hardcopy system. The new system can also adapt to new computing systems and business models quickly as the NCCS adds or upgrades resources or makes other changes that affect resource accounting.

Advanced software technology

Members of the Advanced Software Technology Group (ASTG) have advanced degrees in mathematics, physics, engineering, meteorology, or computer science. This expanded range of expertise is necessary because of the wide range of programming needs of today's NCCS user community. One of the biggest challenges in

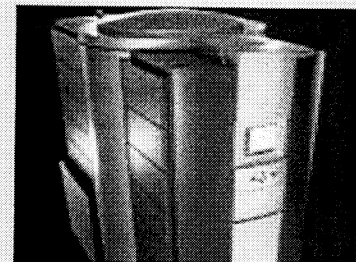
large-scale computing is optimizing the software as major codes have migrated from the traditional vector supercomputing environment to scalable-parallel systems to take advantage of the newer architectures. The expert ASTG staff helps users convert vector codes to parallel codes, thereby improving the algorithm execution to make the best use of the NCCS systems. The ASTG has demonstrated the ability to adapt legacy code to newer algorithms and programming practices in use today.

The ASTG works collaboratively with the users to improve understanding of the requirements and to ensure that code improvements and optimizations are incorporated into production models, allowing users to accomplish more in less time. The NCCS offers one-on-one consulting through the ASTG to help major users optimize codes to run on either the NCCS vector systems or scalable-parallel systems. The goal is to allow scientists to complete their studies more rapidly than before and to increase the size of problems they can tackle. Over the past year, the ASTG has optimized and parallelized

HISTORY

Mid 1990s

The NCCS replaced the Cray Y-MP with a Cray C98. Installation of the Cray C98 marked the first time that the performance of 1 GigaFLOPS per processor was available. In 1995, the NCCS installed its first Cray J90 supercomputer.





atmospheric, ocean, and space science codes and has assisted with the conversion of Fortran 77 code to Fortran 90 to take advantage of the features and benefits of the newer Fortran standard.

In addition to working directly on codes, the ASTG enables users to enhance their skills and improve their own codes through participation in classes taught by the resident experts in the technology. The NCCS tailors its training services to support NCCS users during transitions to new hardware and software systems and to help them use NCCS resources most effectively. NCCS personnel regularly teach short classes on Fortran 90, OpenMP or MPI, and basic optimization. The NCCS has also offered other classes, such as awk and Regular Expressions, in response to the user community's needs. New classes are planned to address new systems, including "bring-your-own-code" classes with vendor experts.

The ASTG started a new initiative to develop a powerful software engineering tool (called *qDoc*) to aid users in the documentation of legacy

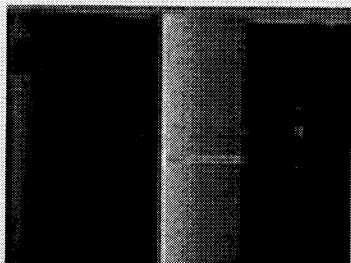
Fortran code through "reverse engineering" technology. The technology will provide the original source code, augmented with comments describing major functionality and made available in hypertext markup language (HTML) format with links to the documentation; a pop-up data dictionary describing the functions of variables, arrays, structures, and other major code objects; and a documentation-linked entity-relationship diagram. The NCCS is developing the *qDoc* tool for initial use on a major meteorological code. After demonstration of the tool's capability, *qDoc* will be made available for broader use.

System support

System support staff serve as administrators for the various machines, keeping them on-line and installing system upgrades when necessary. During the past year, the system administrators have completed several major operating system upgrades and have been actively involved in

HISTORY

Mid 1990s



Even though the J90 processors were less powerful than the C98, the increased number of processors provided substantially more computing capacity. The NCCS mass storage subsystem was the largest such system in the world, with a data storage capacity of 38 TB.



bringing new computing resources on-line in a stable operating configuration.

Another significant function of the systems administrators is maintaining security. The focus of the NCCS security team is on the organizations and projects within the domain of the computing center and of related projects and facilities. The security team is part of a Goddard Center-wide security organization that benefits users and administrators in the NCCS and elsewhere. The NCCS makes its security experts available for consultation to other organizations within Goddard. The security team's chief challenge is finding a way to balance the need to build and maintain effective security with the need to provide public services to the users. An open Goddard network environment is essential to the productivity of the scientific community; however, protecting NCCS computing resources in this type of environment requires significant

effort. The security team is continually evaluating and enhancing the NCCS network topology, proactively monitoring the network, and maintaining intrusion detection efforts.

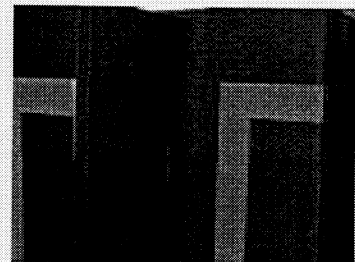
Computer Users Committee

The NCCS constantly strives to address the changing needs of its users and provide the best services and facilities. The NCCS CUC provides a forum for discussion of computer hardware and software, the way the computing and support services are being delivered, and plans for the future. The bimonthly meeting serves as a forum to obtain input from customers, promote communication between the NCCS staff and the user community, announce changes, and work to improve the facilities and their operation. The CUC meetings provide opportunities to address specific issues associated with the NCCS.

The acquisition of the Cray T3E marked the next major shift from vector supercomputing to large-scale parallel computing. By using commodity processors with a high speed interconnect, this system provided the user community with both significant capability and capacity. The T3E was the platform for research in high-performance computing methods.

HISTORY

Late 1990s





Preparing for the Future

The NCCS's commitment to provide exceptional computing resources to the scientific community will remain steadfast. We recognize that dependence on computational power for advances in scientific research is expected to increase dramatically in the 21st century. The NCCS continues to explore effective utilization of both scalable microprocessors and vector supercomputers. The era of the great observatories for both space and Earth sciences will increase the volume of observational data sets by orders of magnitude within the next decade. The NCCS is evaluating new technologies for mass storage to meet these demanding requirements.

As technology continues to change at a rapid pace, we will explore innovative ways to manage these changes effectively for our users. The combined modeling, data analysis, and data assimilation requirements for this era represent a formidable challenge to all aspects of NCCS support. The NCCS is prepared to meet this challenge and will continue its critical role in enabling cutting-edge science research well into the future.

HISTORY

Late 1990s



An SGI Origin 2000 provided a small parallel system with a shared memory architecture that is easier to program than the distributed memory architecture of the T3E. At the end of the 1990s, the J90s were upgraded to Cray SV1s, which returned performance to 1 GigaFLOPS per processor.